Ngoc Tuong Le

GitHub: Link | LinkedIn: Link

Location: District 1, Ho Chi Minh City, Vietnam
Email: lengoctuong23052002@gmail.com | Mobile: +84 34 816 7229

AI Engineer (Intern)

I am a fresh graduate Data Science student focusing strongly on **machine learning and deep learning for computer vision and natural language processing**. I can research, develop, and optimize AI models proven by scientific publications, as well as the ability to process and analyze data. I seek an AI Engineer position to develop my skills further and build a successful career.

Education

University of Science - VNUHCM   District 5, Ho Chi Minh City, Vietnam
Bachelor in Data Science   *Oct 2020 – Oct 2024*

- **Courses**: Probability and Statistics, Linear Algebra, Calculus, Data Structures, OOP, Software Engineer, Database, Data Science, Data Mining, Machine Learning, Big Data, Computer Vision Thesis.
- **GPA**: 3.71.

Achievements

Publication	:	Submitted the paper "CS-MAT: Cross-Shaped Window Mask-Aware Transformer for Large Mask Image Inpainting" to the 2024 17th International Conference on Machine Vision (ICMV 2024).
Award	:	Outstanding Student Science and Technology Research Award at HCMC University of Science
Certificates	:	TOEIC Listening and Reading (Score: 845), Speaking and Writing (Score: 270).
Scholarship	:	Recipient of academic achievement scholarships for several semesters.

Projects

Thesis (Grade: 4/4) - Image Inpainting using generative network combined with Transformer   Oct 2023 – Aug 2024
Technologies Used: PyTorch, NumPy, Matplotlib, OpenCV, scikit-image, scikit-learn, Linux, Git, LaTeX   *GitHub*

- Team size: 2. My responsibilities included research, architecture design, implementation, training, and experiments.
- Developed an innovative approach to reconstruct large missing image areas using a novel Transformer-based generative model.
- Built, trained, and optimized the model on a facial dataset (CelebA-HQ), achieving superior performance compared to several state-of-the-art techniques.
- Application demo video and Demo App. Had the research paper accepted at ICMV 2024, an international conference.

Developing a Text Generation App by fine-tuning GPT   Aug 2023 – Oct 2023
Technologies Used: PyTorch, NumPy, Pandas, Matplotlib, Hugging Face Transformers, Streamlit, LaTeX   *GitHub*

- Team size: 3. I was responsible for researching architectures and fine-tuning the model.
- Researched Transformer and GPT-2 architectures, fine-tuned a pre-trained GPT-2 model on the wikitext-2 dataset, and conducted evaluations.
- Developed and demonstrated a web application for text generation. The model can be found here.

Developing a Machine Learning Pipeline using TensorFlow Extended (TFX)   Feb 2023 – May 2023
Technologies Used: Python, TensorFlow, TFX, Linux, Docker, LaTeX, Trello   *GitHub*

- Team size: 4. I was responsible for researching pipeline components and developing the ML pipeline.
- Researched a sequence of components that implement an ML pipeline, and libraries of tfx.
- Developed a TFX pipeline for a penguin classification model, including Docker-based model serving.

Additional Projects such as Analyzing High School Exam Scores, Analyzing Anime Dataset,... helped me improve my skills, including data wrangling, exploratory data analysis, modeling, and visualization.

Skills and Language

Programming	:	Python (NumPy, Pandas, Matplotlib, SciPy, OpenCV...), C/C++, and SQL/MySQL.
Libraries	:	PyTorch, TensorFlow, PySpark, and sklearn.
Tools	:	Git, Spark, Hadoop, Docker, and cloud (Azure AI/ML).
Data/AI skills	:	Developing, training, and deploying machine learning models; data wrangling, EDA, and visualization.
Deep Learning	:	Knowledge about CNN, GAN, RNN, LSTM, Transformer, Reinforcement Learning,...
Soft Skills	:	Independent research, teamwork, persistence, dedication, and attention to detail.
Language	:	Intermediate proficiency in English; capable of reading and understanding technical documents.